



**University of  
Zurich**<sup>UZH</sup>

**Zurich Open Repository and  
Archive**

University of Zurich  
University Library  
Strickhofstrasse 39  
CH-8057 Zurich  
[www.zora.uzh.ch](http://www.zora.uzh.ch)

---

Year: 2012

---

## **Training of tonal similarity ratings in Non-Musicians: a “Rapid Learning” approach**

Oechslin, Mathias S ; Läge, Damian ; Vitouch, Oliver

**Abstract:** Although cognitive music psychology has a long tradition of expert–novice comparisons, experimental training studies are rare. Studies on the learning progress of trained novices in hearing harmonic relationships are still largely lacking. This paper presents a simple training concept using the example of tone/triad similarity ratings, demonstrating the gradual progress of non-musicians compared to musical experts: In a feedback-based “rapid learning” paradigm, participants had to decide for single tones and chords whether paired sounds matched each other well. Before and after the training sessions, they provided similarity judgments for a complete set of sound pairs. From these similarity matrices, individual relational sound maps, intended to display mental representations, were calculated by means of non-metric multidimensional scaling (NMDS), and were compared to an expert model through procrustean transformation. Approximately half of the novices showed substantial learning success, with some participants even reaching the level of professional musicians. Results speak for a fundamental ability to quickly train an understanding of harmony, show inter-individual differences in learning success, and demonstrate the suitability of the scaling method used for learning research in music and other domains. Results are discussed in the context of the “giftedness” debate.

DOI: <https://doi.org/10.3389/fpsyg.2012.00142>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-98921>

Journal Article

Published Version



The following work is licensed under a Creative Commons: Attribution-NonCommercial 3.0 Unported (CC BY-NC 3.0) License.

Originally published at:

Oechslin, Mathias S; Läge, Damian; Vitouch, Oliver (2012). Training of tonal similarity ratings in Non-Musicians: a “Rapid Learning” approach. *Frontiers in Psychology*:3:142.

DOI: <https://doi.org/10.3389/fpsyg.2012.00142>



# Training of tonal similarity ratings in non-musicians: a “rapid learning” approach

Mathias S. Oechslin<sup>1\*</sup>, Damian Läge<sup>2</sup> and Oliver Vitouch<sup>3</sup>

<sup>1</sup> Faculty of Psychology and Educational Sciences, Geneva Neuroscience Center, University of Geneva, Geneva, Switzerland

<sup>2</sup> Applied Cognitive Psychology, Department of Psychology, University of Zurich, Zurich, Switzerland

<sup>3</sup> Cognitive Psychology Unit, Department of Psychology, University of Klagenfurt, Klagenfurt, Austria

## Edited by:

Christoph T. Weidemann, Swansea University, United Kingdom

## Reviewed by:

Sarah Creel, University of California at San Diego, USA

John Paul Minda, The University of Western Ontario, Canada

## \*Correspondence:

Mathias S. Oechslin, Faculty of Psychology and Educational Sciences, Geneva Neuroscience Center, University of Geneva, Uni-Mail, Bd du Pont-d'Arve 40, 1211 Genève 4, Switzerland.  
e-mail: mathias.oechslin@unige.ch

Although cognitive music psychology has a long tradition of expert–novice comparisons, experimental training studies are rare. Studies on the learning progress of trained novices in hearing harmonic relationships are still largely lacking. This paper presents a simple training concept using the example of tone/triad similarity ratings, demonstrating the gradual progress of non-musicians compared to musical experts: In a feedback-based “rapid learning” paradigm, participants had to decide for single tones and chords whether paired sounds matched each other well. Before and after the training sessions, they provided similarity judgments for a complete set of sound pairs. From these similarity matrices, individual relational sound maps, intended to display mental representations, were calculated by means of non-metric multidimensional scaling (NMDS), and were compared to an expert model through procrustean transformation. Approximately half of the novices showed substantial learning success, with some participants even reaching the level of professional musicians. Results speak for a fundamental ability to quickly train an understanding of harmony, show inter-individual differences in learning success, and demonstrate the suitability of the scaling method used for learning research in music and other domains. Results are discussed in the context of the “giftedness” debate.

**Keywords:** learning, perceptual learning, music, harmony, expertise, mental representations, cognitive maps

## INTRODUCTION

Professional musicians differ from musical novices in their mental and neural representations of tonal relationships (e.g., Bever and Chiarello, 1974; Platt and Racine, 1990; Besson and Faïta, 1995; Andrews et al., 1998; Bigand et al., 1999; Koelsch et al., 1999; Vos and Verkaart, 1999; Müller et al., 2010). In the recent past, the literature has typically either focused on the debate about musical talent and musical skills, or has demonstrated the differences between musicians and novices by means of group comparisons (Schlaug et al., 1995a,b; Pantev et al., 1998; Ohnishi et al., 2001; Shahin et al., 2005; Oechslin et al., 2010a,b). Little research, however, is available on the transition from musical novice to expert. The question of learning progress is obviously not independent of the appropriate tools for measuring such progress. Peoples' auditory percepts are often assessed by verbal descriptions (“What did you hear?” Mikumo, 1992), same-different tasks (Gaab and Schlaug, 2003), or reproduction (vocal or instrumental; Summers et al., 1986; Dalla Bella et al., 2003). Declarative knowledge and specialist vocabulary is a prerequisite for reliable descriptions of what has been heard, and reproduction demands specific motor skills. Musical novices possess little of these. To quantify listening performance, we ideally should assess listeners' responses in a way that is independent of their expertise level.

In this paper we present a measurement procedure based on pairwise similarity ratings (see Materials and Methods) that permits this. We combined this with a short term laboratory training: In a *rapid learning paradigm*, participants had to decide for single

tones and inversions of different major triads whether sound pairs matched each other well. Participants received immediate feedback, derived from experts' average judgments, after the rating of each pair. The changes in the cognitive maps computed from these ratings provide quantitative insight into the learning process.

An alternative approach based on cognitive priming has been established in the past to test listeners' harmonic understanding of musical material. This paradigm usually introduces a harmonic context, followed by unexpected musical events to be detected (Maess et al., 2001; Tillmann et al., 2006). These target stimuli challenge the listener's expectation, which depends on the individual's implicit and explicit knowledge on harmonic progressions and musical syntax. Non-musicians can reach considerable rating accuracy in musical priming paradigms, pointing to a remarkable implicit proficiency (Koelsch et al., 2000). However, presenting subtle violations of harmonic expectation (e.g., finishing a cadence by first inversion instead of the root position) revealed that harmonic target detection is considerably modulated by the degree of musical expertise, as visible on both behavioral performance and brain activity levels (James et al., 2008).

Creating an experimental paradigm based on quantifiable ratings of tone/chord pairs opens a complementary window onto the cognitive organization of the harmonic space. Participants respond on a numeric scale, providing a quantitative basis for calculating individual and groupwise mental representations. These cognitive maps of harmonic relations can then be compared as a function of musical expertise (see Materials and Methods

subsection). This elemental approach to studying cognitive representations beyond expertise boundaries is not limited to the field of music, but can be applied to every perceptual modality in order to analyze any categorically organized field of knowledge.

The field of music is an excellent example for the long-term coexistence of different and partly irreconcilable theoretical accounts for individual learning potential. The next two subsections give a brief juxtaposition of the major players in this debate, namely talent vs. expertise accounts.

### THE TALENT CONUNDRUM

The existence of musical talent has long been a matter of vigorous debate. As with the assumption of talents for other domains, core questions concern the nature, frequency, and role of a special giftedness for making and processing music. Since the notion of talent essentially implies innateness, the ultimate goal would be to identify genetic components that permit an early prognosis of who is musical (and who is not). Sir Francis Galton, the early promoter of the idea of “hereditary genius” (Galton, 1869), used family records of famous composers to demonstrate heritability as early as in the nineteenth century.

Today, despite genetics experiencing something of a heyday, an influential school of expertise theorists holds that inborn talent is a hopelessly ill-structured concept from folk psychology – a notion that is culturally prevalent in Western societies, but is ultimately circular and non-explanatory (Howe et al., 1998). Howe et al. (1998) argue that even if it does fundamentally exist, talent – defined as a rare, innate, domain-specific ability to excel or to progress with extraordinary ease in a certain area – is scientifically useless as long as no valid prediction can be made about future achievements. At the core of the debate is the question whether practice is a *necessary* condition (which few researchers would object to) or a *sufficient* condition for exceptional performance, the latter idea being the position of the most hardcore variant of expertise theory (see Vitouch, 2005, for a critical overview). As a reaction to this position, several methodological suggestions have been made in order to adequately test the impact of individual factors (including biological ones), such as applying testing-the-limits designs grounded in life-span development research to heterogeneous samples (Baltes, 1998) or conducting longitudinal studies following “talented” and “untalented” subjects, both practicing (Vitouch, 1998; see also other commentaries to Howe et al., 1998).

### MUSICAL EXPERTISE

Prognostic validity has always been a tough challenge for theories and psychometric instruments of giftedness, and particularly so in the musical domain in which beginning and success are typically separated by decades. Expertise research, on the other hand, has demonstrated that the cumulative amount of deliberate practice (Ericsson et al., 1993) is an excellent predictor for professional achievement, and has put forward the “10-year rule of necessary preparation” (Simon and Chase, 1973) to account for expert performance from a training-oriented perspective (for the case of musical expertise, see overviews by Sloboda, 1991; Vitouch, 2005; Lehmann and Gruber, 2006). The question remains, however, whether a long-term regimen of deliberate musical practice is a sufficient condition to achieve musical mastery, or whether

there is a necessary hidden component that separates the budding young Mozarts from the tone-deaf. The question is evidently of high societal, political, and practical relevance: Which, and how many, individuals should be given sponsored access to music education; on what grounds and with what expectations of success? In other words: How much musical expertise can the average junior citizen achieve?

### THE ROLE OF ECOLOGICAL VALIDITY

Some authors even argue that music reception is such a common daily habit in our society that practically everybody quickly becomes a “private expert” in this area (Sloboda, 1991; Smith, 1997). At the very least, it seems evident that “receptive expertise” does not automatically require “productive expertise” (which is commonplace in many music ethnicities, but comparatively rare in our tradition) – otherwise, this would be bad news for professional critics. While “musical illiteracy,” the inability to read and write music, is generally widespread, several studies have demonstrated that musical novices strongly improve on listening tasks as soon as ecologically valid instructions and formats are used (Smith et al., 1994). For instance, in an interval recognition task, if a major sixth is difficult to categorize, the task becomes much easier by simply renaming it the *My Way* interval, which opens access to an adequate mental representation in a quick, reliable and effortless manner. Novices can perform intriguingly well in such auditory tasks, and in many others in the domains of relative pitch, perception of dissonance (Koelsch et al., 2000), recognition of musical phrasing (Palmer et al., 2001), recognition of tunes based on infinitesimal excerpts of music (Schellenberg et al., 1999), and so forth (for an overview, see Bigand and Poulin-Charronnat, 2006).

An example of the intricacy of the ecological validity issue comes from an experimental study by Platt and Racine (1990). The problem is that groups with different musical background may base their musical judgments on different sources (e.g., implicit judgments vs. theory-guided judgments). Platt and Racine (1990, Exp. 1) compared musicians and non-musicians with regard to perceptual similarity judgments of single tones and triads (with the single tones also being constituent of the respective triad). For instance, what is “more similar,” (1) c’ and a C major triad c’/e’/g’, or (2) g’ and the same triad? The authors seem to assume that all participants made their judgments based on perceptual impressions of similarity, and report structured differences between the groups’ response patterns. We, in contrast, argue that due to the deliberately vague instructions<sup>1</sup>, musicians typically fell back on their theoretical knowledge, and decided that a chord should be conceptually “most similar” to either its tonic or to the melody tone (“root trackers” vs. “melody trackers”; Platt and Racine, 1990, p. 418). Non-musicians had no such theoretical knowledge available, and as the entire experiment was performed without feedback (since the experimenters could not tell which answer would be “correct” either), they produced a clutter of responses without a

<sup>1</sup>“The subject’s task was to push one of two buttons to indicate which of the component notes sounded ‘most similar’ to the triad. [...] If the subjects asked questions about the basis for the similarity judgment, they were told this was one of the questions that the experimenters were trying to answer” (Platt and Racine, 1990, p. 417).

discernible pattern. According to our interpretation, the meaning of the entire task was quite alien to this group.

### THE RAPID LEARNING PARADIGM

The general question of this paper is: Are musically “untalented” (i.e., not particularly talented) people able to quickly achieve specific competencies of music perception by means of simple training methods? We are thinking of average listeners who have no systematic record of practical music education (such as instrument lessons). They form the target population for our rapid learning paradigm: Assume that a defined musical listening task, which initially cannot be adequately tackled, can be easily and rapidly trained, with the result of expert-like post-training performance. It would then seem legitimate to argue that (a) learning plays a major role for this task, and (b) the task is broadly accessible, and not restricted to a talented minority. If this can be shown for a multitude of relative pitch class tasks, it demonstrates that (a) and (b) probably hold for relative pitch *per se*.

### Perceptual Learning

In many instances of such learning processes, *perceptual learning* (Goldstone, 1998) plays a major role. Perceptual learning has the defining property that at the beginning of a learning or categorization process, naïve learners cannot even understand *what is to be learned*. Not only are they unable to categorize correctly; they do not perceive any distinct categories at all. In the process of feedback-based perceptual learning, implicit categories emerge which enable us to see (or hear) the world with new eyes (or ears). For instance, using computer-morphed faces, Goldstone and Steyvers (2001) showed that after roughly 200 trials, participants had learned to dichotomously classify such faces according to a latent bipolar dimension, with different dimensions “perceptually taught” to different experimental groups. These dimensions are discovered and constructed in the course of the learning process. Absolute pitch seems to be an intriguing case of perceptual learning, though limited to a sensitive period for acquisition up to age 6 (Takeuchi and Hulse, 1993; Ward, 1999; Russo et al., 2003; Vitouch, 2003; Deutsch et al., 2006). We assume that also in most listening tasks, novices merely lack the perceptual categories that are a prerequisite for dealing with these tasks appropriately. New cortical cell assemblies must connect as the neural substrate of these perceptual abilities. This highlights that we are dealing with domain-specific questions of neural plasticity (and its limitations) here, but through a behavioral lens of observation.

### Pairwise ratings of tonal material

Pairwise ratings have been used in earlier studies. Hubbard’s (1998) participants did chord-chord comparisons, while Platt and Racine (1990) and Hubbard and Datterri (2001) examined tone-chord ratings. Krumhansl and Kessler (1982) used so-called probe tone ratings to investigate tonal structures in Western music. Their participants had to indicate how well a target tone matches a previously established key (introduced as a rising diatonic scale), using all chromatic tones as targets. These and similar procedures serve as models for both our diagnosis paradigm (pairwise similarity ratings) and the match/do not match ratings in the rapid learning phase.

### TONAL REPRESENTATION OBSERVED BY MULTIDIMENSIONAL SCALING

Let us finally review some earlier examples for relational modeling of sounds through multidimensional scaling (MDS). Krumhansl et al. (1982) asked their participants to rate chord pairs on a scale from 1 (follows poorly) to 7 (follows well) and thus assessed the individually perceived quality of the succession. All major triads were compared that can be formed on steps I–VII of the keys C major and F-sharp major ( $=2 \times 7$  triads). An evaluation by MDS (cf. Materials and Methods section) revealed that both keys span a two-dimensional space, in which they are distributed in a completely consistent manner across two corresponding zones. This solution corresponds to a representation of the circle of fifths reduced to the minimum – represented only by two keys. Krumhansl later compiled a large number of further investigations, which can be seen as a type of ontology of the circle of fifths representation (Krumhansl, 1990, 1991). Based on these findings we assume that (i) musically trained listeners are able to cluster tonic and inversions of the same key due to their tonal similarity and that (ii) clusters of different keys are clearly geometrically separated but spatially organized according to the circle of fifths.

### AIM AND HYPOTHESES

With regard to the conception of the rapid learning paradigm, we assume that novices possess all necessary abilities to perceive acoustic material in a naïve sense. The objective therefore consists in *extending* certain skills of relative (in contrast to absolute) pitch. The novices are confronted with conditions that enable perceptual learning in the sense of a purely perception-based conveying of learned contents as elements that can be clearly categorized. Using suitable stimuli, novices should be influenced so that certain mental representations are formed – but without any confrontation with declarative knowledge. This type of knowledge transmission and implicit knowledge consolidation therefore yields a reorganization of existing patterns without having something explicitly presented to the learner.

The general aim of this study is to determine if rapid learning is possible in the domain of tonal relationships. We fundamentally assume that simple feedback-based training will be reflected in the results even in the very short term. Participants are expected to have their mental “sound maps” consistently moving toward the average map of the expert model, with the ideal case of a post-training novice map being indistinguishable from an expert map. Although some participants may not profit from the training due to a large number of potential reasons (from motivation to individual differences to training duration), we expect a major share of our sample to make rapid progress. If this central assumption holds, it would put the results from a number of earlier studies (such as Platt and Racine, 1990) into a new perspective.

### MATERIALS AND METHODS

#### PARTICIPANTS

Twenty-three musical novices from Switzerland participated in this study, and were randomly assigned to an experimental group with 16 persons and a control group with 7 persons. The experimental group consisted of 6 women and 10 men with an average age of 28 years (range 23–38 years). Twelve of them never played a

musical instrument, and the other 4 had begun to learn an instrument in childhood but never managed to go beyond beginner's level. In addition, nobody of them had played any musical instrument during the last 15 years, and none of the participants was educated in harmonics.

The control group consisted of two women and five men with an average age of 26 years (range 24–27 years) that had the same non-musical history as the experimental group. They were asked at two separate appointments to perform the task of similarity ratings. By comparing the maps of these two assessments we were able to evaluate whether repeated similarity ratings alone, without feedback, lead to an improvement in listening performance.

### NON-METRIC MULTIDIMENSIONAL SCALING AND PROCRUSTEAN TRANSFORMATION

“Cognitive maps” allow to assess and compare individual knowledge structures, and learning-related reorganization of knowledge in a certain field (Läge et al., 2008). These maps display the internal representation of aspects or sections of the physical or social environment, which is developed by one's experiences and subjective interests (Marx and Läge, 1995). This method refers to the concept of geographical representation originally introduced by Tolman (1948) which provides a mapping of the individually perceived similarity or dissimilarity between elements in our environment (Läge et al., 2008). In the present study we utilized NMDS to get access into the structural organization of knowledge about Western tonal harmonics represented by two subsets of major tonalities. Via NMDS the revealed matrix of pairwise similarity ratings (here on a nine-point scale) can be transformed into a cognitive knowledge map (Marx and Heij, 1989; Borg and Groenen, 1997; Läge, 2001; Streule et al., 2006). This geometrical representation visualizes the subject's perceived similarity between objects, as larger or smaller distances, allowing an interpretation of semantic relations apparent in a specific field of knowledge. Cognitive maps are widely accepted as an appropriate model for those cases of factual knowledge in which certain objects (which can be described by a set of features) exist side by side (Läge et al., 2008). To assess the quality of knowledge we compared each learner map with the expert map using procrustean transformation. A procrustean transformation (Gower and Duksterhuis, 2004) superimposes two maps on top of each other by scaling, shifting, rotating, and mirroring the configuration of objects in order to achieve maximal congruence (Läge et al., 2008). Importantly, aiming for maximal congruence this procedure may change map-sizes but not the geometrical relations manifested by object positions forming the maps. The information concerning the distances reflecting this comparison can be expressed numerically: the resulting divergence of two maps is determined as the Object-Loss between two corresponding objects and as the AverageLoss ( $\text{AvgLoss} = \text{mean ObjectLoss}$ ) between two cognitive maps (Läge, 2001).

### OPERATIONALIZATION OF THE RAPID LEARNING PARADIGM

In step 1 (baseline assessment), the ability of a novice to correctly match tonics and inversions of the same chord is determined. This occurs through pairwise similarity judgments (nine-point scale) between chords and single tones from a pool of 16 sounds.

Based on the individual matrix of all similarity judgments, an individual “sound map” is calculated through NMDS. By means of procrustean transformation, each individual sound map can be compared with an expert model, and the deviations between maps can be quantified. The expert model (see **Figure 1**), empirically based on averaged proximities across experts, has been evaluated in a pilot study (Oechslin et al., 2006). We investigated 13 musicians (mean instrumental experience = 30 years), most of them being conservatory teachers, by applying exactly the same similarity rating exercise as introduced in the present study.

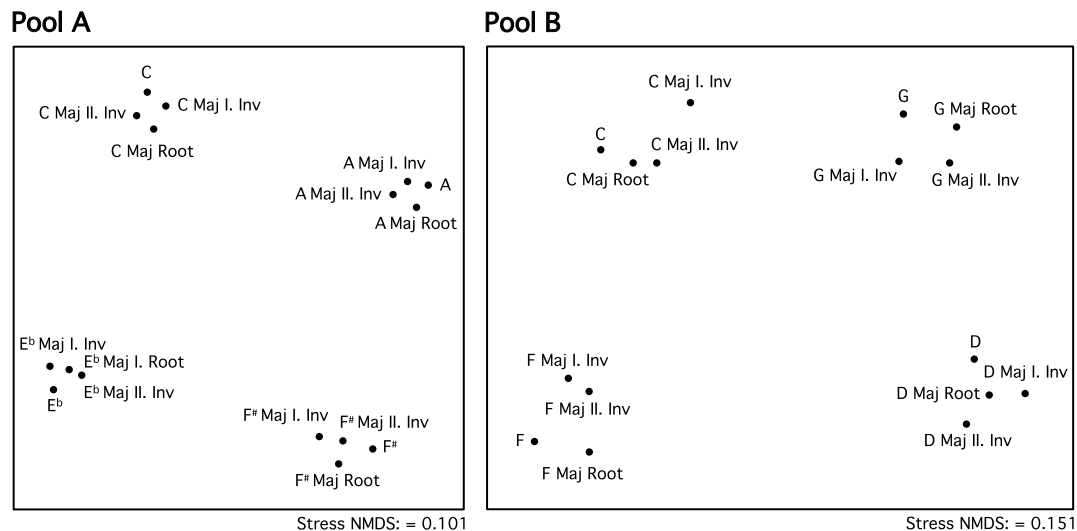
In step 2 (the rapid learning phase), a feedback-driven training is carried out. Novices were presented with chords and single tones on a pairwise basis. Listening to these sound pairs, the subjects were asked to indicate *whether* they match well or not. (Consider this difference to the non-dichotomous judgments required in steps 1 and 3.) The training begins with clearly distinguishable pairs and increases in difficulty as soon as a person masters a trained level (see Materials and Methods section for details).

In step 3 (post-training assessment), a knowledge diagnosis is performed exactly as in step 1. Depending on structure and consistency, any qualitative improvement of the individual “sound map,” relative to the experts' map, reflects a progress in listening performance.

### STIMULI AND ASSESSMENT OF TONAL SIMILARITY RATINGS

We implemented two different pools (A/B) of tonalities: both pools included four different tonalities, each represented by three major triads based on the fundamental chord [in root position (root), first inversion (I. Inv) and second inversion (II. Inv)] and corresponding tonics (as single tones). Thus each of the pools consists of 16 sounds that represent the following tonalities: *Pool A*: C major, A major, F-sharp major, and E-flat major. With respect to the circle of fifths these four keys are found at minor third and tritone (diminished fifth) distance to one other and therefore represent maximum between-tonality distances, and thus highest dissimilarity by means of musical functions in terms of this harmonic model. *Pool B*, by contrast, includes neighboring tonalities within the circle of fifths: D major, G major, C major, and F major. These keys therefore represent minimum between-tonality distances in the circle of fifths, and thus highest similarity by means of musical function.

All chords and tones were generated on a Roland 5600s synthesizer (*Piano* mode) and post-processed with the softwares *Wave* and *CoolEdit.Pro* resulting in exact length of 1500 ms and normalized amplitudes to avoid uncontrolled stimulus salience (resolution: 16 bit (stereo), sampling rate: 44.1 kHz). For each pool we included all possible sound pairs (tone-tone, chord-chord, tone-chord), resulting in  $n \cdot (n - 1) / 2 = 120$  pairs in total. The positions of pairs' two tonal constituents are determined by the following principle: The position of chords and single tones were defined by aiming for the smallest possible tonal interval between the lowest and the highest tone calculated among all tones contributing to the tonal pair. This voicing rule results in a minimal effort concerning mental octave transpositions and thus provides an optimal setting for non-musicians to handle paired sounds in a naïve manner. For instance, in chord/single tone pairs (of the



**FIGURE 1 | “Expert maps Pool A and Pool B.”** This figure depicts the mean expert representations of Pool A (left) and Pool B (right), computed by averaging the individual ( $n = 13$ ) NMDS maps (Oechslin et al., 2006) which are based on the same 240 similarity ratings (120 for each pool) as performed by musical novices (see Materials and Methods). These maps provide a basis for

expert vs. novice comparisons by means of evaluating the AvgLoss (=average deviation of a certain novice map from the expert map) for each individual, which thus represents the spatial discrepancy between the individual novice and the expert solution at certain points of time (abbreviations: root = root position of a major chord; 1/2. Inv = first/second inversion of a major chord).

same key), the single tone is always identical (same octave position) to its complement embedded in the chord. Consequently chord/chord pairs within the same tonality always share two of their three tones.

In an earlier experiment (Oechslin et al., 2006), the same 120 tonal pairs were presented to 13 professional musicians (30 years of experience on average) in order to assess pairwise similarity ratings on a nine-point scale (9 = extremely similar, 1 = extremely dissimilar). The revealed complete triangular matrices of similarity values were averaged across all 13 musical experts. The averaged matrix of similarity rating for each pool was subjected to a two-dimensional NMDS. The two resulting mean expert maps are presented in **Figure 1**. Henceforth these two maps were used in the present study as an expert model for the quality assessment of musical novices’ listening performances (see procrustean transformation).

Both expert maps indicate that four sounds associated to one of the four tonalities are displayed adjacent and form sharply delimited clusters. The spatial arrangement of the clusters corresponds the circle of fifths and for that reason providing an explanation in terms of harmonics that appears adequately: The clusters of Pool B (neighboring tonalities) reveal a semi-circle due to maximal inter-cluster similarity; while the clusters of Pool A (opposite tonalities) drift apart toward the corners of a square, due to maximal inter-cluster dissimilarity. Regarding the map of Pool B we like to point out that single tones (tonics) are slightly shifted in direction (left) of that very tonality they occur as the fifth in the chord’s root position. We consider this as strong evidence for a high resolution of harmonic decoding mechanisms in professional musicians, and moreover, from a rather methodical point of view, as an indication for the robustness of the applied NMDS algorithm (ROBUSCAL) regarding the individual consistency of the data.

As mentioned above, experts’ similarity ratings on these two sound pools serve as expert models and further as a basis for the development of the “rapid learning paradigm.” Due to the greater dissimilarity of chords in Pool A, it should be easier, compared to Pool B, for musical novices to distinguish between sounds of the same than of different tonalities, since Pool B includes neighboring keys with maximum similarities. We thus expect that the revealed cognitive maps reflect the participants’ increased challenge for separating and clustering the tonalities of Pool B compared to Pool A.

The present experiment consists of two basic components: (i) the measure of the cognitive structures of tonal relations (based on tonal similarity ratings) and (ii) a training unit in order to improve this. The measurement of the cognitive structuring occurs through pairwise similarity judgments of two sequentially played sounds (chords or single tones) on a nine-point similarity scale. They were presented in individual random orders. Using NMDS, the matrix of similarity judgments provided by a person was transformed into a geometrical solution, a two-dimensional Euclidean map. Performing procrustean transformation, each individual map was then placed onto the corresponding expert map (of Pool A or B). The extent of AvgLoss indicates the deviation of the tonal representation from the expert model. Furthermore, individual matrices of similarity ratings were subjected to a hierarchical cluster analysis (mean value model). Based on hypothetical four-cluster solutions (due to the expert models), we evaluated categorization errors based on the similarity judgments of each participant at two different time points ( $t_0$  vs.  $t_2$ ).

In sum, the AvgLoss (comparison of the whole structure) and the number of categorization errors (discrimination of same/different key dichotomy) served as parameters for quantifying individual learning progress.



## THE “RAPID LEARNING” PARADIGM

Here the same sound pairs were presented as the ones used for the similarity rating task. Participants had to accomplish different levels of difficulty (examples in **Figure 2**): the task was to decide whether the sound pairs matched or not. The participants scored, indicated by a smiley, if the yes/no response corresponded the harmonic norm (given by expert maps). The goal was to collect seven smiles with the restriction that mistakes would nullify the collected scores at a certain level. Therefore the corresponding difficulty level starts over. After seven *consecutive* correct answers, the level was accomplished and participants moved up to the next level.

At the first level, we presented those sound pairs which were rated by highest similarity and highest dissimilarity with respect to the expert models. In contrast, the fourth level contained sound pairs with similarity values that were adjacent to the mean value due to the expert's ratings and thus represent the most difficult and ambiguous exercise.

In order to allocate each sound pair to one of the four levels, we ranked them according to their mean similarity given by experts. The ranking-order consisting of 120 pairs was divided into 8 equal units (=15 pairs per unit), so that in each case two units were merged and defined as one level (=30 pairs): For Level 1, the first and the eighth unit were taken (including the highest and lowest similarities), for Level 2, the second and the seventh unit, for Level 3 the third and the sixth unit, and for Level 4 the fourth and the fifth unit. Following this approach the task (good/bad match) becomes more difficult from level to level. During training

sessions, participants passed through all four levels. Pool A and B were trained at separate appointments.

























The complete experiment spanned five sessions and was organized as following: at the first appointment, participants were presented with  $2 \times 120$  sound pairs for similarity ratings. At the second to fifth appointment, one of the two pools was trained with the procedure described above. Directly after each training session participants were asked to give similarity ratings on the trained pool. Musical novices were trained twice with each Pool in one of these four orders: AABB, BBAA, BAAB, ABBA. The following abbreviations will be further used in the *Results* section for the similarity ratings at different points in time:  $t_0$  = baseline,  $t_1$  = after first training,  $t_2$  = after second training.

## RESULTS

### TONAL SIMILARITY RATINGS

Each participant of the control group provided four NMDS maps: one of each for Pool A and Pool B at each appointment. These 28 maps ( $=7$  subjects  $\times$  2 Pools  $\times$  2 appointments) were fitted onto the expert models of Pool A and B separately by performing via procrustean transformation. For every individual map, we thus calculated an AvgLoss that expresses the difference with respect to the expert map. To investigate the main effect of similarity ratings we statistically tested the 14 maps from the first appointment (mean AvgLoss = 0.67, SD = 0.13) against the 14 maps of the second appointment (mean AvgLoss = 0.71, SD = 0.11). A paired  $t$ -test (two-tailed) revealed no significant improvement of similarity ratings between these two timepoints ( $t_{13} = -1.504$ ,  $p = 0.156$ ).

Examples: Rapid Learning Paradigm

	Level	Sound Pair	Positions	Mean expert similarity rating	Correct answer (y/n)
Pool A	1	 E <sup>1</sup> -Maj root		8.69	y
		 A-Maj root		3.31	n
	2	 C-Maj I.Inv		7.69	y
		 C-Maj II.Inv		3.62	n
	3	 E <sup>1</sup> -Maj II.Inv		5.85	y
		 E <sup>1</sup> -Maj root		4.15	n
	4	 A-Maj II.Inv		5.23	y
		 E <sup>1</sup> -Maj I.Inv		4.54	n
Pool B	1	 D-Maj root		8.69	y
		 D-Maj I.Inv		3.85	n
	2	 C-Maj II.Inv		8.15	y
		 F-Maj II.Inv		4.23	n
	3	 D-Maj I.Inv		7.61	y
		 C-Maj I.Inv		5.00	n
	4	 C-Maj root		6.92	y
		 F-Maj I.Inv		5.85	n

**FIGURE 2 | “Examples: rapid learning paradigm.”** Here we show representative examples of sound pairs (sequential order of sounds is arbitrary) presented in the rapid learning paradigm, for both pools (A/B) and each level (1–4) of difficulty. The positions within the expert maps are schematically displayed: dots correspond to clusters of each tonality, the position of sounds correspond to the white dots. In case of only one white dot, both sounds belong to the same tonality. For every level we

show one example of relatively high [correct answer: yes (y)] and one of relatively low similarity [correct answer: no (n)], which represent expert's ratings given as mean similarity values. Mean expert ratings provide the base for the allocation of each sound pair to one of the four levels to be accomplished by performing the rapid learning paradigm (abbreviations: root = root position of a major chord; I./II. Inv = first/second inversion of a major chord).

Moreover, the number of categorization errors (determined by hierarchical cluster analysis; for details see above) did not decrease significantly either: the number of categorization errors corresponded an average of 8.1 (SD 2.1) in the first measurement and of 7.8 (SD 1.9) in the second. A paired  $t$ -test (two-tailed) revealed no significant difference ( $t_{13} = 0.717$ ,  $p = 0.486$ ). Taken together, this indicates that systematic improvements in the experiment can be attributed to the training.

In order to illustrate these improvements, **Figure 3** shows the training history of two representative good learners from the experimental group, with Pool A on the left and Pool B on the right side (baseline:  $t_0$ , after first training:  $t_1$ , after second training:  $t_2$ ). The NMDS solution for the similarity ratings of the learner (large dots) was placed onto the expert model using procrustean transformation (small dots). For illustration purposes corresponding dots are linked to one another. The subject on the left side started with a strong organization of Pool A, however, still improved; the subject on the right side started with a weak organization of Pool B and achieved a reasonably well structure comparable with the quality of the baseline map of the other subject in terms of AvgLoss indicating notabene the degree of deviation of two compared maps. Accordingly, the baseline map ( $t_0$ ) of the left hand subject in **Figure 3** (AvgLoss: 0.31) considerably resembles the expert model. Therefore, even before training, this learner perceived the sounds from the same key as most similar to each other, and the keys that are in a tritone relationship as most dissimilar (i.e., keys that are located opposite to each other on the map: C–F $^\sharp$  and A–E $^b$ ). However, the four inter-cluster distances (i.e., between sounds of the same tonality) are far not as closely associated as in the expert solution: the four large dots of each key are much further apart than the corresponding small dots. In the second ( $t_1$ ) and third ( $t_2$ ) map (below), the clusters manifest systematically, and the AvgLosses thus decreases from 0.15 to 0.09. The latter value of 0.09 is the lowest of the whole experiment; therefore, among musical novices, this very map approximates the expert model most closely.

Initially, we expected Pool B (with the neighboring keys in the circle of fifths) to appear substantially more difficult compared to Pool A. This is markedly demonstrated by the right hand subject (**Figure 3**), since the AvgLoss of 0.81 in the baseline map is considerably high. This participant was not able to correctly allocate the sounds to each other. Following the first training session ( $t_1$ ), the map clearly improves and the AvgLoss decreases to 0.47: The four sounds of the two keys (F-/D-maj), located at the edge of the selected section from the circle of fifths, can be clearly detected as clustered. In contrast, the sounds of the two keys in between (C-/G-maj) do not seem to be clustered. At least, following the second training session ( $t_2$ ), the participant had further improved his harmonic representation (AvgLoss: 0.38), however the G major cluster still appears to be distributed from top to bottom in the map.

## QUANTIFYING TRAINING PROGRESS

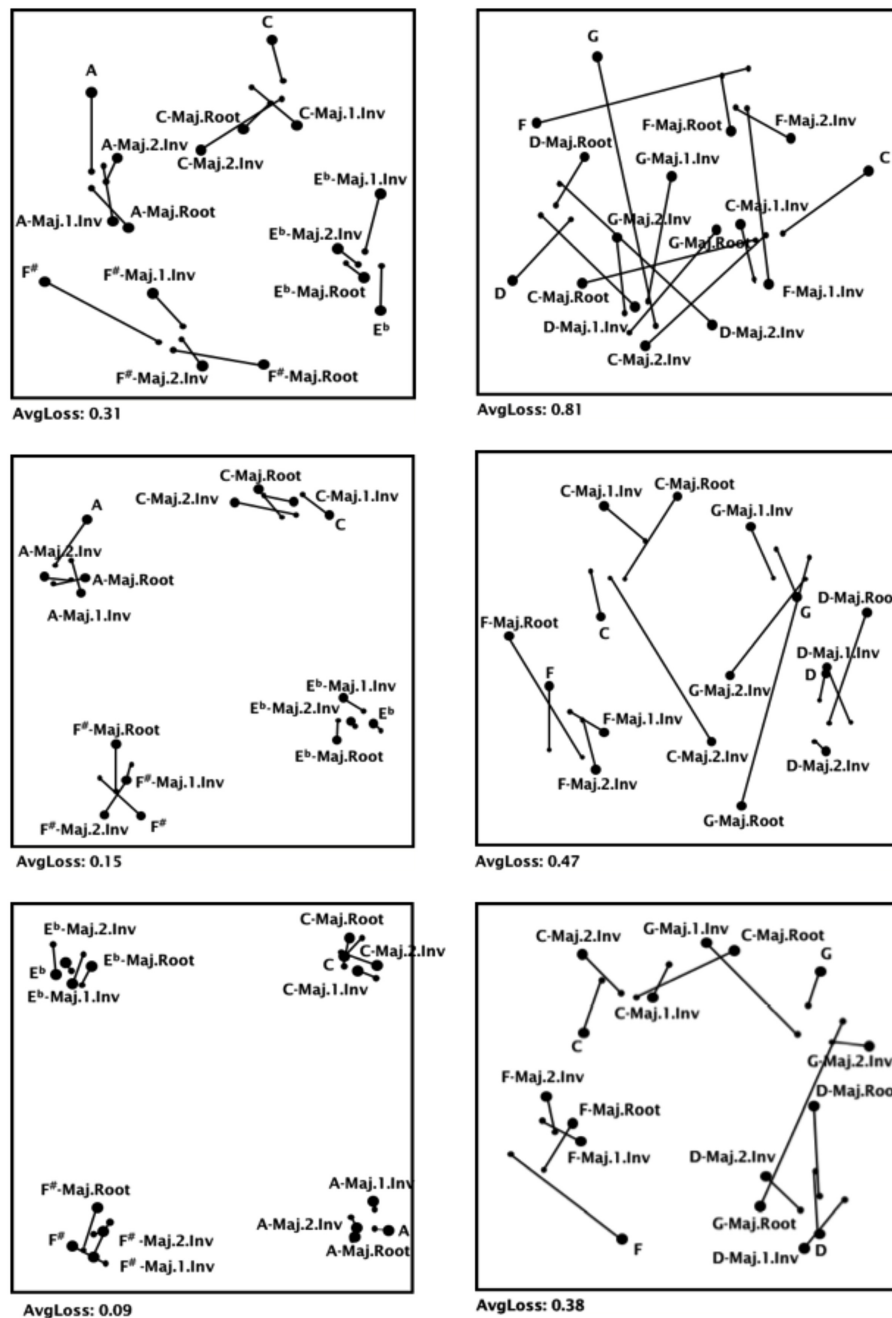
In the following paragraph we aim to expose how to statistically quantify the training progress by taking into account the full experimental sample. As shown before, two NMDS maps can be quantitatively described through the AvgLoss in terms of their similarity. To understand the comparison of samples, the sound

elements can be replaced by the information structure, the similarity matrix, provided by one single subject. By laying all pairs of available maps and performing procrustean transformations, a dissimilarity matrix results – containing values that reflect the dissimilarity of every possible pair of maps. Since AvgLoss reflects for the average geometrical distance between all objects (here: sounds) of two maps, we speak of ObjectLoss when focusing on a distance between two objects (here: individual learners). The aforementioned dissimilarity matrix is again scaled using NMDS and thus reveals a so-called loss oriented meta-map (LOMM, see Läge, 2001), in which each point incorporates a complete individual NMDS map (based on the individual similarity ratings). More precisely, LOMMs represent the relational position of the subjects based on the similarity of their individual cognitive maps (Egli et al., 2008). **Figure 4** shows the two LOMMs for Pool A (left) and Pool B (right). Dots represent NMDS maps of a learner both before training (green dots) and after the second training session (red dots). In addition, this Figure also takes into account all individual expert maps (black dots) from the earlier study (Oechslin et al., 2006) clustered around the expert model (AvgExp).

Individual experts' maps strongly cluster in the right hand part of the maps (with one exception in Pool B). Dots representing the maps of trained non-musicians broadly constitute the middle and left area of the LOMMs. In addition, for illustration purposes, two semi-circles imply the interpretative differentiation between good and moderate maps (inner semi-circle) and between moderate and poor maps (outer semi-circle). An arrow pointing toward the expert model indicates that the respective person improved his representation with respect the expert model and has thus benefited from the training. If an arrow however does not approach the expert model or even diverges, the training has thus not improved the performance of this person. The complete picture provides a good impression of the heterogeneity of the results: some learners improve their cognitive map very strongly after the two training sessions, while others do not improve at all. Some individuals, whose maps were fairly good from the outset, even manage to enter the tight cluster of the maps of the professional musicians. Apparently, these participants whose maps were poor at the outset, by contrast, none of them managed such a great leap. Here, too, some participants benefited substantially from the training, but not everybody.

In order to examine more precisely which participants benefited from the training and which did not, we suggested two different evaluations: Firstly, we tested for significant improvements of maps by subjecting ObjectLosses to paired  $t$ -tests. For this purpose, deviations of individual objects in NMDS maps were evaluated by comparing them with expert models by procrustean transformations (Egli et al., 2008; Läge, 2001). In the case of a single map (examples **Figure 3**) the AvgLoss represents the mean of 16 ObjectLosses (one for each sound object), each reflecting the distance between a sound object's position in the learners' map and in the expert solution. This therefore enabled us to test for each subject whether the influence of training yields a significant improvement relative to the expert model. Accordingly, we performed  $t$ -tests for repeated measures ( $n = 16$  sound objects) and compared for each participant the ObjectLosses of  $t_0$  against  $t_2$ . The analysis revealed (at  $\alpha = 0.05$ ) that





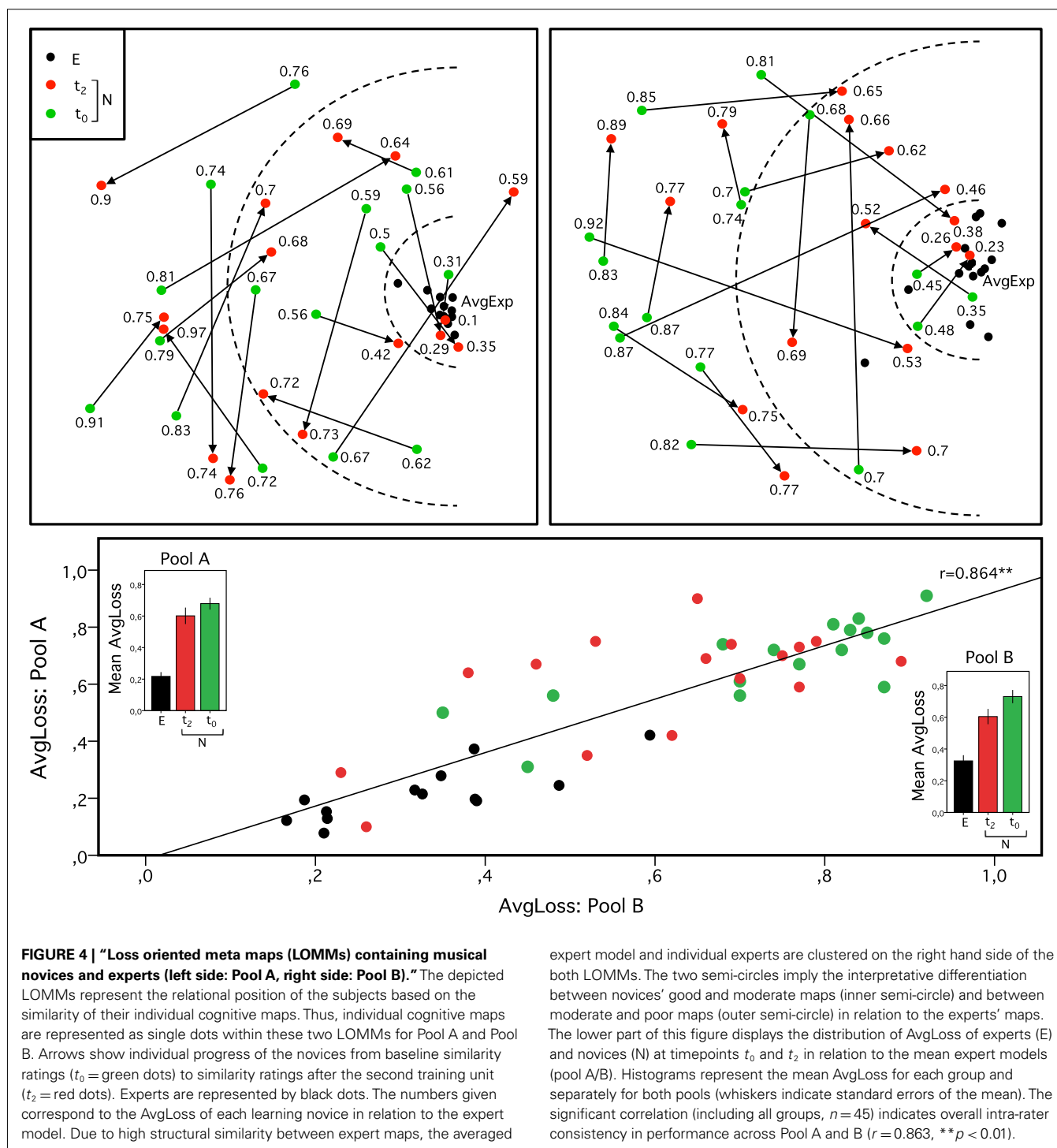
**FIGURE 3 | “Two individual learning histories.”** This figure delineates Procrustean transformations of two learning novices at three successive time points: first row: baseline ( $t_0$ ), second row: after the first training ( $t_1$ ) and third row: after the second training ( $t_2$ ) for Pool A (left column) and Pool B (right column). The individual maps (big dots) are placed over the averaged expert

maps (small dots); connecting lines mark the corresponding deviations. For both representative “learning histories”, the AvgLoss (from top to bottom) is constantly decreasing; thus the novice maps consistently converge with the average expert maps. (Abbreviations: root = root position of a major chord; 1/2. Inv = first/second inversion of a major chord.)

7 out of 16 participants improved in Pool A following the two training sessions (participants 1, 3, 4, 8, 10, 13, 14), and 8 persons improved in Pool B following the two training sessions (1, 3, 4, 5, 6, 8, 9, 14).

Secondly, in order to validate these results, we additionally conducted the following person-based evaluation on the group level:

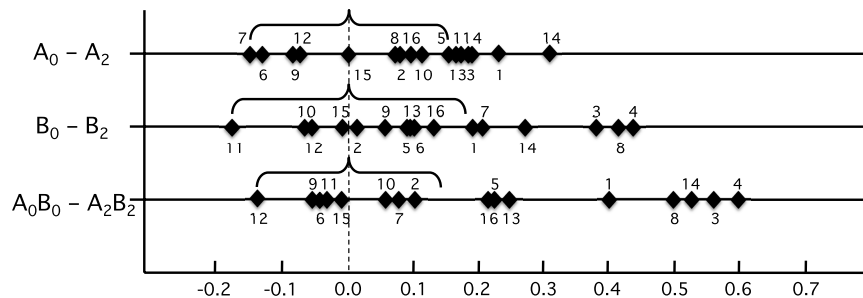
There is a convincing reason (namely the training) why the listening performance of a person improves in the course of the experiment. However, there is no reason why the listening performance should become systematically worse through the experiment (represented by negative AvgLosses in **Figure 5**). Deteriorations in the LOMM map can therefore be considered randomly distributed. In



particular, we assume that they define one side of a random distribution around a null-effect of training. In this case, the other side is described best through the symmetry assumption, which predicts that learners will improve or deteriorate to approximately the same degree (indicated by brackets in **Figure 5**). Concerning the subjects whose maps improved beyond the area of deteriorations, one can assume a systematic training effect. **Figure 5** illustrates the

improvements and deteriorations of the individual maps, measured as changes in distance in the LOMM map to the expert model. Here, too, it is plausible to assume a systematic improvement in performance after the two training sessions of about half the sample.

Approximately half of the participants significantly improved their listening performance, which led to a significant overall



**FIGURE 5 | “Distribution of individual improvements in terms of AvgLoss.”** This figure depicts the individual training effects (numbers 1–16 correspond to the participants) by subtracting the AvgLosses after the second training ( $t_2 = A_2, B_2, A_2B_2$ ) from AvgLosses before training ( $t_0 = A_0, B_0, A_0B_0$ ). We calculated these effects separately for Pool A (upper row), Pool B (middle row), and based on averaged values of Pool

A and B (bottom row). Brackets embrace subjects that show an unsystematic pattern of learning; based on the assumption that subjects form a random distribution around the null-effect of training. Subjects characterized by highly positive AvgLoss differences (to the right of the brackets) provide evidence for a systematic effect of training.

**Table 1 | “Improvements of AvgLoss in musical novices.”**

Mean AvgLoss of musical novices					
Pool A			t-Test		
Comparison	$t_0$	$t_1$	$t_2$	t-Value	p-Value
$t_0$ vs. $t_1$	0.68	0.63	–	1.839*	<0.05
$t_1$ vs. $t_2$	–	0.63	0.60	0.556	0.29
$t_0$ vs. $t_2$	0.68	–	0.60	2.429*	<0.05
Pool B			t-Test		
Comparison	$t_0$	$t_1$	$t_2$	t-Value	p-Value
$t_0$ vs. $t_1$	0.73	0.66	–	1.831*	<0.05
$t_1$ vs. $t_2$	–	0.66	0.61	0.691	0.25
$t_0$ vs. $t_2$	0.73	–	0.61	1.742◇	0.05

This table reports the improvements of musical novices comparing AvgLosses due to similarity ratings at three points in time ( $t_0$  = baseline;  $t_1$  = after first rapid learning phase;  $t_2$  = after second rapid learning phase). The three rows within each section are in favor of all possible comparisons between the three points in time representing the learning history. Paired t-tests (one-tailed) indicate whether subjects improved their performance in relation to the expert maps. Asterisks indicate the level of significance [t-value (\*):  $p < 0.05$ , t-value (◇):  $p = 0.05$ ].

training effect (see Table 1, Comparisons  $t_0/t_2$ ). At the same time, results in Table 1 reveal that the learning effect is primarily accomplished in the first training session (i.e., between  $t_0$  and  $t_1$ ). Here, we found significant differences for both Pool A and Pool B. The average improvement in the second training session is only half as large and not significant neither for Pool A nor for Pool B. Accordingly, we conclude that the implemented training provided a short term effect, however, did not lead to a linear improvement over time. Pool B, incidentally, seemed to be more difficult for the participants prior to the training (mean AvgLoss of 0.73 for Pool B compared to 0.68 for Pool A). After the training sessions, this difference disappeared completely (mean AvgLoss of 0.61 for Pool B compared to 0.60 for Pool A).

So far this evaluation reveals a general improvement of the learners’ tonal representation maps. This can be attributed to the following perspectives: (a) the formation of four clusters that enclose the four sounds of representing each key; and, presumably much more demanding, (b) the correct allocation of these clusters in terms of the circle of fifths. The AvgLoss due to procrustean transformation behaves according to criteria of a good listening performance, represented by accurate similarity judgments. A hierarchical cluster analysis, by contrast, would be sensitive for the clustering of the categories (yet without assuming spatial relations between categories). For this reason, we further present an evaluation of the similarity ratings by cluster analysis. For each individual novice’s similarity matrix, a four-cluster solution is calculated (HCL, mean value model). The minimum number of displacements is calculated that is required to produce a perfect four-cluster solution constituting the four tonalities. By within-subjects comparisons, using paired t-tests, we evaluated whether non-musicians’ performances improve, by means of a decrease of category errors, as a consequence of training. Expert models of Pool A/B revealed flawless categorizations of sounds according to their tonality (Table 2 presents the mean values of the permutations). They decrease in Pool A from 8.50 (before training) through 6.87 to 6.44 after training session 2, and in Pool B from 8.69 through 7.44 to 7.44 (training session 2 therefore had no effect for Pool B). Akin to AvgLosses, the analyses of the category errors let suggest that the training effect is primarily attributed to the first training session.

An individual evaluation in structural analogy to Figure 4 results in seven persons with significant learning effects in Pool A or Pool B (those subjects who also showed learning success in the evaluation with the AvgLosses). The aspect of pure chord categorization is therefore well in line with a consideration of the complete structure.

## DISCUSSION

Linking the “rapid learning” paradigm to a measurement format independent of expertise, we could show that approximately half of the musical novices in our sample substantially improve their listening performance (regarding Western harmonic relations)

**Table 2 | “Improvements of tonal categorization in musical novices.”**

Mean category error of musical novices					
Pool A			t-Test		
Comparison	t <sub>0</sub>	t <sub>1</sub>	t <sub>2</sub>	t-Value	p-Value
t <sub>0</sub> vs. t <sub>1</sub>	8.50	6.87	–	3.499**	<0.01
t <sub>1</sub> vs. t <sub>2</sub>	–	6.87	6.44	0.590	0.28
t <sub>0</sub> vs. t <sub>2</sub>	8.50	–	6.44	2.792**	<0.01
Pool B			t-Test		
Comparison	t <sub>0</sub>	t <sub>1</sub>	t <sub>2</sub>	t-Value	p-Value
t <sub>0</sub> vs. t <sub>1</sub>	8.69	7.44	–	2.660**	<0.01
t <sub>1</sub> vs. t <sub>2</sub>	–	7.44	7.44	0.000	0.50
t <sub>0</sub> vs. t <sub>2</sub>	8.69	–	7.44	3.024**	<0.01

This table reports the improvements of musical novices comparing mean category errors, evaluated by cluster analysis due to similarity ratings of three points in time (t<sub>0</sub> = baseline; t<sub>1</sub> = after first rapid learning phase; t<sub>2</sub> = after second rapid learning phase). The three rows within each section are in favor of all possible comparisons between the three points in time representing the learning history. Paired t-tests (one-tailed) indicate whether subjects improved their performance by means of category errors. Asterisks indicate the level of significance [t-value (\*\*): p < 0.01].

through a simple, feedback-based training. The LOMMs show that several individuals advanced very quickly toward the performance level of professional musicians.

The implications of these results are threefold: First, the listening performance can apparently be trained using the simplest of means and without a great investment of time. We therefore conclude that we are dealing with a phenomenon of cognitive organization of “knowledge,” which must already be present in novices at least in an implicit manner. Second, this improvement in performance was only evident for half of the sample. As we were not able to identify any motivational factors for this (all participants made a real effort to accomplish the task), an inter-individual variability of either the cognitive structuring performance or the listening performance (or both) seems to be responsible. The stability of these inter-individual differences, however, could only be determined using temporally extended retest data, and additional (educational) experimental treatments. Differences due to cognitive strategies, for instance, could be characterized by high volatility. Third, the chosen format of similarity judgments and of cognitive maps proved to be well suited for measuring directly comparable music perception data independently of the music theoretical degree of expertise of the participants.

Let us begin a more detailed consideration of these three conclusions with the latter, methodological, aspect. In the introduction, we pointed out that the question “What did you hear?” was frequently asked using a format that presupposes explicit music theoretical knowledge or practical musical skills. The similarity judgments approach proved to be neutral in this respect. Similar to other areas of expertise (Läge, 2001; Läge et al., 2005; Schlatter and Läge, 2005; Egli et al., 2006), it enabled both novices and experts to provide responses that are appropriate to their

knowledge and could be differentiated sufficiently for both. The modeling by means of NMDS and the distance measurement through procrustean transformations allows for direct and meaningful comparisons beyond expertise boundaries. Consequently, the path chosen in this work can be recommended as a methodological tool for future music psychological studies, and also for other domains of learning progress. The “rapid learning” paradigm also did its bit: It led to a quick improvement in the classification and differentiation of sounds, at least with half of the participants. This shows the fundamental trainability in this field, and indeed with a relatively low expenditure.

The training approach chosen might appear to some to be a resurrection of conditioning experiments. We would not necessarily contradict such an impression as the learning does indeed occur on a level on which there is no need for declarative support. But that was precisely the aim of the experiment: to examine whether a quick success that is easy to induce can also arise without the declarative imparting of music theoretical knowledge in adults. This question can be answered with a clear “yes.”

Of course, as our approach was performance-oriented, we cannot make conclusive statements about the underlying cognitive processes established, or the solution strategies used by the trained novices. It is always conceivable that the same observable output is achieved by very different means (cf. the “zombie argument” in consciousness research, or “Morgan’s cannon” in animal behavior and animal cognition research). If pigeons can learn to discriminate between pictures with vs. without humans shown (Herrnstein and Loveland, 1964) or between music from Bach vs. Stravinsky (Porter and Neuringer, 1984) by means of simple operant conditioning, why should similar processes not also account for the learning success of our subjects? This question about the actual processes of perceptual learning underlying our participants’ new implicit knowledge cannot be adequately answered from our data. While the “sameness,” or structural similarity, of the underlying mental representations and of the perceptual cues that novices are using relative to the experts cannot be proven in our setting, we could demonstrate strong correspondence for some trained novices on the observable level of their tonal similarity ratings.

The improvement in performance turned out to be heterogeneous across participants: Some improve strongly (and even reach the level of the professional musicians), while others remain at the initial level. Although the latter also accomplish all training levels, nothing changes in their mental structuring. For the interpretation of this result, it can be assumed that the feedback-based training has in principle already taken effect (in order to successfully complete all training levels solely through random answers, a participant would have required much more rounds than were observable in the experiment). Nevertheless, a perfect and reliable handling of the task material is not necessary to achieve the required number of correct answers: In order to complete a level in time, it is perfectly adequate if a person is correct in 70% or 80% of cases – at some point she will be sufficiently lucky then and achieve the necessary series of hits. In this respect, it cannot necessarily be concluded from the experiment that all participants had really mastered the given task.

In the context of the giftedness debate described in the introduction, it is revealing to see that a very large range was found in

the resulting cognitive maps. Several persons – who some would probably describe as “gifted” – improve their similarity judgments so tremendously that they can behaviorally keep up with musical experts. The mere rate of successful learners, however, leads the talent concept into a paradox here: If giftedness suddenly appears for about 50% of musically inactive adults, the defining property of rarity (cf. Howe et al., 1998) is violated to an extent which challenges the entire concept. Of course, one could speculate that the number of achievers would proportionally decrease as listening tasks become more and more demanding (a testing-the-limits approach; Lindenberger and Baltes, 1995; Baltes, 1998). Still, the remarkable learning success of many of our naïve participants seems to be at odds with the expectations of the traditional talent account in music. There obviously is a danger of theory immunization if one could just declare every task that a large number of “untalented” people can accomplish with some training as “too easy” after the fact. This means that both the talent camp and the expertise camp have to increase the precision of their *a priori* predictions of what can be acquired under which circumstances (see below), in order to allow paradigmatically relevant empirical tests of trainability effects.

As a wrap-up of the effects with the “rapid learning” paradigm, we can say that about half of the novices behaved similarly to how experts do as a result of the learning phase (especially profiting from the first of the two training sessions). They perceptually achieved a cognitive model of tonal relationships resembling the standard model of musicians. This means that tenet (a), “learning plays a major role for this task” (see the outline in the Introduction, The Rapid Learning Paradigm), is corroborated, and tenet (b), “the task is broadly accessible and not restricted to a talented minority,” holds for at least 50% of the sample. Note that this has been achieved by the simplest of means: No lessons, no explanations, no teachers; just performance-based feedback in a very restricted time frame. In contrast to the findings reported by Platt and Racine (1990), the here introduced feedback approach thus enabled us to communicate to non-musicians *how* sound pairs have to be assessed by means of similarity – namely by putting the focus on root tone relations. Perceptual learning can therefore be considered an important strategy to mediate harmonic knowledge. Additional importance of our contribution is related to the analysis set-up: Cluster analysis and NMDS allowed to study in detail the re-/organization of harmonic knowledge and to quantify learning progress on the group and subject level.

## REFERENCES

- Andrews, M. W., Dowling, W. J., Halpern, A. R., and Bartlett, J. C. (1998). Identification of speeded and slowed familiar melodies by younger, middle-aged, and older musicians and nonmusicians. *Psychol. Aging* 13, 462–471.
- Aufegger, L., and Vitouch, O. (in press). Tales of talent: Rapid learning of acoustic instrument recognition. To appear in *Proceedings of the 12th International Conference on Music Perception and Cognition, ICMPC 2012*.
- Baltes, P. B. (1998). Testing the limits of the ontogenetic sources of talent and excellence. *Behav. Brain Sci.* 21, 407–408.
- Besson, M., and Faïta, F. (1995). An event-related potential (ERP) study of musical expectancy: comparison of musicians with non-musicians. *J. Exp. Psychol. Hum. Percept. Perform.* 21, 1278–1296.
- Bever, T. G., and Chiarello, R. J. (1974). Cerebral dominance in musicians and nonmusicians. *Science* 185, 537–539.
- Bigand, E., Madurell, F., Tillmann, B., and Pineau, M. (1999). Effect of global structure and temporal organization on chord processing. *J. Exp. Psychol. Hum. Percept. Perform.* 25, 184–197.
- Bigand, E., and Poulin-Charronnat, B. (2006). Are we “experienced listeners?” A review of the musical capacities that do not depend on formal musical training. *Cognition* 100, 100–130.
- Bigand, E., Tillmann, B., Poulin-Charronnat, B., and Mandler, D. (2005). Repetition priming: is music special? *Q. J. Exp. Psychol.* 58A, 1347–1375.
- Borg, I., and Groenen, P. (1997). *Modern Multidimensional Scaling: Theory and Applications*. New York: Springer.

The study introduced here has some obvious limitations with respect to the following questions: What exactly is learned (a question due to implicit knowledge transfer), at which age, with which personal characteristics and life trajectories; and how quickly, easily, reliably, and stably? Further research is inevitable to elucidate especially the sustainability of the learning effects demonstrated here. It would not be surprising to find that long-term stability is as low as short term plasticity is high: Performance will certainly depend on the opportunity and will to regularly deal with pertinent musical material.

The fact that *only* about 50% of our participants benefit from the rapid learning paradigm questions that improvements are due to sensory priming effects based on repetitive presentation (rapid learning paradigm). This is supported by previous research observing effects of repetition priming: In contrast to words, pictures, and environmental sounds, chord processing was not facilitated by repetition (Bigand et al., 2005). Different harmonic functions appeared to be more influential on the performance than the repetition of tonal targets. This finding speaks in favor of musical novices using higher cognitive functions to evaluate tonal targets in musical context. Thus our results let us assume that a cognitive representation of harmonic relationships is either already stored but “sleeps” until we present the determining (here implicit) cues, or can be rapidly acquired at least in some cases in a gradual process of perceptual learning. However, to this end it remains unclear which individual factors are critical for an accurate performance relative to expertise-like cognitive processing of Western tonal harmonics.

In conclusion, we predict that rapid learning can be demonstrated with similar ease in other related domains, such as interval, chord, or timbre identification (preliminary evidence for the latter is presented by Aufegger and Vitouch, in press). While the acquisition or development of absolute pitch seems to be firmly restricted to a critical age, relative pitch tasks seem to be well-trainable in a large portion of musically naïve adults. This also makes sense in the light of evolutionary accounts to music (Wallin et al., 2000; Vitouch and Ladinig, 2009): While a certain individual variation of musical aptitude is not surprising, it is the general “ability for music” of the species *H. sapiens* that broadly stands in the foreground.

## ACKNOWLEDGMENTS

We are indebted to Lars Rogenmoser for valuable comments on an earlier version of the paper.

- Dalla Bella, S., Peretz, I., and Aronoff, N. (2003). Time course of melody recognition: a gating paradigm study. *Percept. Psychophys.* 65, 1019–1028.
- Deutsch, D., Henthorn, T., Marvin, E., and Xu, H.-S. (2006). Absolute pitch among American and Chinese conservatory students: prevalence differences, and evidence for a speech-related critical period. *J. Acoust. Soc. Am.* 119, 719–722.
- Egli, S., Schlatter, K., Streule, R., and Läge, D. (2006). A structure-based expert model of the ICD-10 mental disorders. *Psychopathology* 39, 1–9.
- Egli, S., Streule, R., and Läge, D. (2008). The structure-based expert model of the mental disorders – a validation study. *Psychopathology* 41, 286–293.
- Ericsson, K. A., Krampe, R. T., and Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychol. Rev.* 100, 363–406.
- Gaab, N., and Schlaug, G. (2003). The effect of musicianship on pitch memory in performance matched groups. *Neuroreport* 14, 2291–2295.
- Galton, F. (1869). *Hereditary Genius: An Inquiry into its Laws and Consequences*. London: Macmillan.
- Goldstone, R. L. (1998). Perceptual learning. *Annu. Rev. Psychol.* 49, 585–612.
- Goldstone, R. L., and Steyvers, M. (2001). The sensitization and differentiation of dimensions during category learning. *J. Exp. Psychol. Gen.* 130, 116–139.
- Gower, J. C., and Duksterhuis, G. B. (2004). *Procrustes Problems*. Oxford: Oxford University Press.
- Herrnstein, R. J., and Loveland, D. H. (1964). Complex visual concept in the pigeon. *Science* 146, 549–551.
- Howe, M. J. A., Davidson, J. W., and Sloboda, J. A. (1998). Innate talents: reality or myth? *Behav. Brain Sci.* 21, 399–442.
- Hubbard, T. L. (1998). Listeners can discriminate among major chord positions. *Percept. Mot. Skills* 87, 891–897.
- Hubbard, T. L., and Datter, D. L. (2001). Recognizing the component tones of a major chord. *Am. J. Psychol.* 114, 569–589.
- James, C. E., Britz, J., Vuilleumier, P., Hauert, C.-A., and Michel, C. M. (2008). Early neuronal responses in right limbic structures mediate harmony incongruity processing in musical experts. *Neuroimage* 42, 1597–1608.
- Koelsch, S., Gunter, T. C., Friederici, A. D., and Schröger, E. (2000). Brain indices of music processing: “non-musicians” are musical. *J. Cogn. Neurosci.* 12, 520–541.
- Koelsch, S., Schröger, E., and Ter-vaniet, M. (1999). Superior attentive and pre-attentive auditory processing in musicians. *Neuroreport* 10, 1309–1313.
- Krumhansl, C. L. (1990). *Cognitive Foundations of Musical Pitch*. New York: Oxford University Press.
- Krumhansl, C. L. (1991). Music psychology: tonal structures in perception and memory. *Annu. Rev. Psychol.* 42, 277–303.
- Krumhansl, C. L., Bharucha, J., and Castellano, M. A. (1982). Key distance effects on perceived harmonic structure in music. *Percept. Psychophys.* 32, 96–108.
- Krumhansl, C. L., and Kessler, E. J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychol. Rev.* 89, 334–368.
- Läge, D. (2001). *Ähnlichkeitsbasierte Diagnostik von Sachwissen [Similarity-based diagnostics of declarative knowledge]*. Unpublished habilitation thesis, University of Zurich, Zurich.
- Läge, D., Daub, S., Bosia, L., Jäger, C., and Ryf, S. (2005). *Die Behandlung ausreißer-behafteter Datensätze in der Nonmetrischen Multidimensionalen Skalierung – Relevanz, Problemanalyse und Lösungsvorschlag [The treatment of samples with outliers in non-metric multidimensional scaling]*. AKZ Research Report No. 21. Zurich: Angewandte Kognitionsspsychologie, University of Zurich.
- Läge, D., Oberholzer, R., Egli, S., and Streule, R. (2008). Assimilative learning with the aid of cognitive maps. *Int. J. Emerg. Technol. Learn.* 3, 29–33.
- Lehmann, A. C., and Gruber, H. (2006). “Music,” in *The Cambridge Handbook of Expertise and Expert Performance*, eds K. A. Ericsson, N. Charness, P. J. Feltovich, and R. R. Hoffman (Cambridge: Cambridge University Press), 457–470.
- Lindenberger, U., and Baltes, P. B. (1995). Testing-the-limits and experimental simulation: two methods to explicate the role of learning in development. *Hum. Dev.* 38, 349–360.
- Maess, B., Koelsch, S., Gunter, T. C., and Friederici, A. D. (2001). Musical syntax is processed in Broca’s area: an MEG study. *Nat. Neurosci.* 4, 540–545.
- Marx, W., and Hejj, A. (1989). *Subjektive Strukturen [Subjective structures]*. Göttingen: Hogrefe.
- Marx, W., and Läge, D. (1995). *Der Ideologische Ring [The Ideological Ring]*. Göttingen: Hogrefe.
- Mikumo, M. (1992). Encoding strategies for tonal and atonal melodies. *Music Percept.* 10, 73–81.
- Müller, M., Höfel, L., Brattico, E., and Jacobsen, T. (2010). Aesthetic judgments of music in experts and laypersons – an ERP study. *Int. J. Psychophysiol.* 76, 40–51.
- Oechslin, M. S., Imfeld, A., Loenneker, T., Meyer, M., and Jäncke, L. (2010a). The plasticity of the superior longitudinal fasciculus as a function of musical expertise: a diffusion tensor imaging study. *Front. Hum. Neurosci.* 3:76. doi:10.3389/neuro.09.076.2009
- Oechslin, M. S., Meyer, M., and Jäncke, L. (2010b). Absolute pitch – functional evidence of speech-relevant auditory acuity. *Cereb. Cortex* 20, 447–455.
- Oechslin, M. S., Vitouch, O., and Läge, D. (2006). “Von Quintenzirkel und Kleinterzproblem – die kognitive Repräsentation von Dur-Dreiklängen bei Musikern [The cognitive representation of major triads in musicians],” in *Perspektiven Psychologischer Forschung in Österreich. Proceedings zur 7. Wissenschaftlichen Tagung der Österreichischen Gesellschaft für Psychologie*, eds B. Gula, R. Alexandrowicz, S. Strauß, E. Brunner, B. Jenull-Schiefer, and O. Vitouch (Lengerich: Pabst), 85–91.
- Ohnishi, T., Matsuda, H., Asada, T., Aruga, M., Hirakata, M., Nishikawa, M., Katoh, A., and Imabavashi, E. (2001). Functional anatomy of musical perception in musicians. *Cereb. Cortex* 11, 754–760.
- Palmer, C., Jungers, M. K., and Jusczyk, P. W. (2001). Episodic memory for musical prosody. *J. Mem. Lang.* 45, 526–545.
- Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L.-E., and Hoke, M. (1998). Increased auditory cortical representation in musicians. *Nature* 392, 811–814.
- Platt, J. R., and Racine, R. J. (1990). Perceived pitch class of isolated musical triads. *J. Exp. Psychol. Hum. Percept. Perform.* 16, 415–428.
- Porter, D., and Neuringer, A. (1984). Music discriminations by pigeons. *J. Exp. Psychol. Anim. Behav. Process.* 10, 138–148.
- Russo, F. A., Windell, D. L., and Cuddy, L. L. (2003). Learning the “special note”: evidence for a critical period for absolute pitch acquisition. *Music Percept.* 21, 119–127.
- Schellenberg, E. G., Iverson, P., and McKinnon, M. C. (1999). Name that tune: Identifying popular recordings from brief excerpts. *Psychon. Bull. Rev.* 6, 641–646.
- Schlatter, K., and Läge, D. (2005). *Lernen mit Wissensstrukturkarten [Learning with knowledge structure maps]*. AKZ Research Report No. 26. Zurich: Angewandte Kognitionsspsychologie, University of Zurich.
- Schlaug, G., Jäncke, L., Huang, Y., and Steinmetz, H. (1995a). In vivo evidence of structural brain asymmetry in musicians. *Science*, 267, 699–701.
- Schlaug, G., Jäncke, L., Huang, Y., Staiger, J.-F., and Steinmetz, H. (1995b). Increased corpus callosum size in musicians. *Neuropsychologia* 33, 1047–1055.
- Shahin, A., Roberts, L. E., Pantev, C., Trainor, L. J., and Ross, B. (2005). Modulation of P2 auditory-evoked responses by the spectral complexity of musical sounds. *Neuroreport* 16, 1781–1785.
- Simon, H. A., and Chase, W. G. (1973). Skill in chess. *Am. Sci.* 61, 394–403.
- Sloboda, J. A. (1991). “Musical expertise,” in *Toward a General Theory of Expertise: Prospects and Limits*, eds K. A. Ericsson and J. Smith (Cambridge: Cambridge University Press), 153–171.
- Smith, J. D. (1997). The place of musical novices in music science. *Music Percept.* 14, 227–262.
- Smith, J. D., Kemler Nelson, D. G., Grohskopf, L. A., and Appleton, T. (1994). What child is this? What interval was that? Familiar tunes and music perception in novice listeners. *Cognition* 52, 23–54.
- Streule, R., Egli, S., Oberholzer, S., and Läge, D. (2006). “Adaptivity in e-learning – provided by cognitive maps,” in *EDEN 2006 Annual Conference: E-Competences for Life, Employment and Innovation*, eds A. Szűcs and I. Bø (Budapest: EDEN), 16–21.
- Summers, J. J., Hawkins, S. R., and Mayers, H. (1986). Imitation and production of interval ratios. *Percept. Psychophys.* 39, 437–444.
- Takeuchi, A. H., and Hulse, S. H. (1993). Absolute pitch. *Psychol. Bull.* 113, 345–361.
- Tillmann, B., Koelsch, S., Escoffier, N., Bigand, E., Lalitte, P., Friederici, A. D., and von Cramon, D. Y. (2006). Cognitive priming in sung and instrumental music: activation of inferior frontal cortex. *Neuroimage* 31, 1771–1782.
- Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychol. Rev.* 55, 189–208.

- Vitouch, O. (1998). The rage to drink, or: frontiers of expertise. *Behav. Brain Sci.* 21, 428–429.
- Vitouch, O. (2003). Absolutist models of absolute pitch are absolutely misleading. *Music Percept.* 21, 111–117.
- Vitouch, O. (2005). “Erwerb musikalischer expertise [acquisition of musical expertise],” in *Allgemeine Musikpsychologie (Enzyklopädie der Psychologie, Vol. D/VII/1)*, eds T. H. Stoffer and R. Oerter (Göttingen: Hogrefe), 657–715.
- Vitouch, O., and Ladinig, O. (eds). (2009). Music and evolution. *Music. Sci.* 2009–2010 (Special Issue).
- Vos, P. G., and Verkaart, P. P. (1999). Inference of mode in melodies. *Music Percept.* 17, 223–239.
- Wallin, N. L., Merker, B., and Brown, S. (eds). (2000). *The Origins of Music*. Cambridge, MA: MIT Press.
- Ward, W. D. (1999). “Absolute pitch,” in *The Psychology of Music*, 2nd Edn, ed. D. Deutsch (San Diego, CA: Academic Press), 265–298.
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Received: 23 January 2012; accepted: 21 April 2012; published online: 17 May 2012.
- Citation: Oechslin MS, Läge D and Vitouch O (2012) Training of tonal similarity ratings in non-musicians: a “rapid learning” approach. *Front. Psychology* 3:142. doi: 10.3389/fpsyg.2012.00142
- This article was submitted to *Frontiers in Cognitive Science*, a specialty of *Frontiers in Psychology*.
- Copyright © 2012 Oechslin, Läge and Vitouch. This is an open-access article distributed under the terms of the Creative Commons Attribution Non Commercial License, which permits non-commercial use, distribution, and reproduction in other forums, provided the original authors and source are credited.